

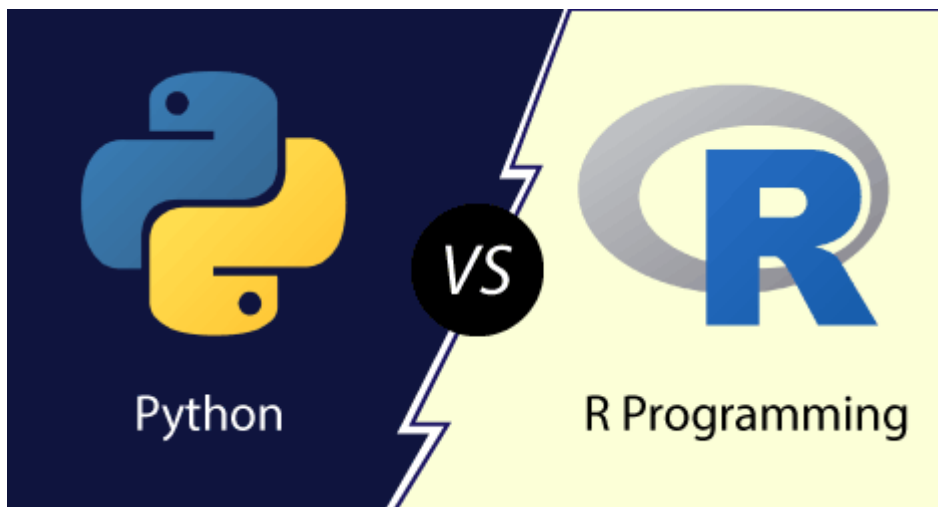
R v/s Python

R

Academics and statisticians have developed R from the past two decades. R now has one of the richest ecosystems for data analysis. In CRAN (open-source repository), there are about 12000 packages available. For whatever analysis we want to do, it is possible to find a library. The rich diversity of the library makes R the first choice for statistical analysis and especially for specialized analytical work.

Python

Python can do a lot of the same functions as R: data wrangling, apps, engineering, feature selection web scraping, and so on. Python is a tool for deploying and implementing large-scale machine learning. Python code is easy to maintain and stronger than R. In the past few years, Python did not have enough data analysis and machine learning libraries. But recently, Python is catching up and providing advance APIs for machine learning or artificial intelligence. Most data science jobs can be done with five Python libraries: Pandas, Numpy, Scikit-learn, Scipy, and Seaborn.



R is considered as the best programming language for any statistician because it contains a comprehensive list of statistical and graphical methods. On the other hand, Python can function similar to R, but data scientists or data analysts prefer it due to its simplicity and high performance. R is a powerful scripting language and is highly flexible with a vibrant community and a resource bank, while Python is a widely used, object-oriented language that is easy to learn and debug.

Let's start making the comparison between R and Python-based on some factors such as speed, flexibility, and popularity, etc.

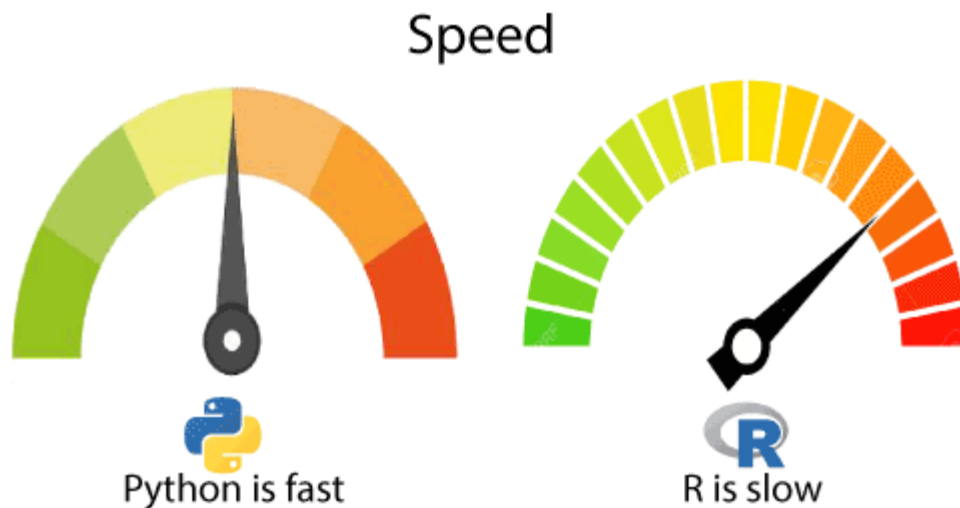
Ease of Learning

R is used for statistical analysis of data. R has a steep learning curve. The people who have less or no experience in programming will find it difficult in the beginning. R is not so hard, once we get a grip of the language. On the other hand, Python emphasizes productivity and code readability, which make it one of the simplest programming languages. It is better due to the ease of learning and understanding.



Speed

Because of the low-level programming language, R requires longer codes for simple procedures. A long code takes much time in execution, and is the reason for the reduced speed. Python is a high-level programming language which is preferred for building critical, yet fast applications. Python codes are too simple and reduce in length, which takes less time to run.



Data Handling Capabilities

R is convenient for analysis due to a large number of packages, the advantage of using formulas, and easily usable tests. It can be used for basic data analysis without any installation of packages. The Python package was an issue for data analysis but has been improved with recent versions. Numpy and Pandas packages are used for data analysis in Python. It is also suitable for parallel calculations. From the discussion, it is clear that in data handling, both are good.



Graphics and Visualization

We can efficiently and more effectively understand the visualized data rather than the raw data. In R, several packages provide advanced graphical capabilities. Visualizations

are important when choosing data analysis software. Python has some amazing visualization libraries, but they are complex and give a clear output. R advanced graphical capabilities make R stronger than Python.

Graphics & Visualization



Looser



Winner

Flexibility

When we compare R and Python with a flexibility factor, then we find that both the languages are flexible. Because in R, it is easy to use complex formulas, and R statistical tests and models are also available, and we can use them easily. Python is flexible when it comes to build something from scratch. Python is also used for scripting a website and other applications.

Flexibility



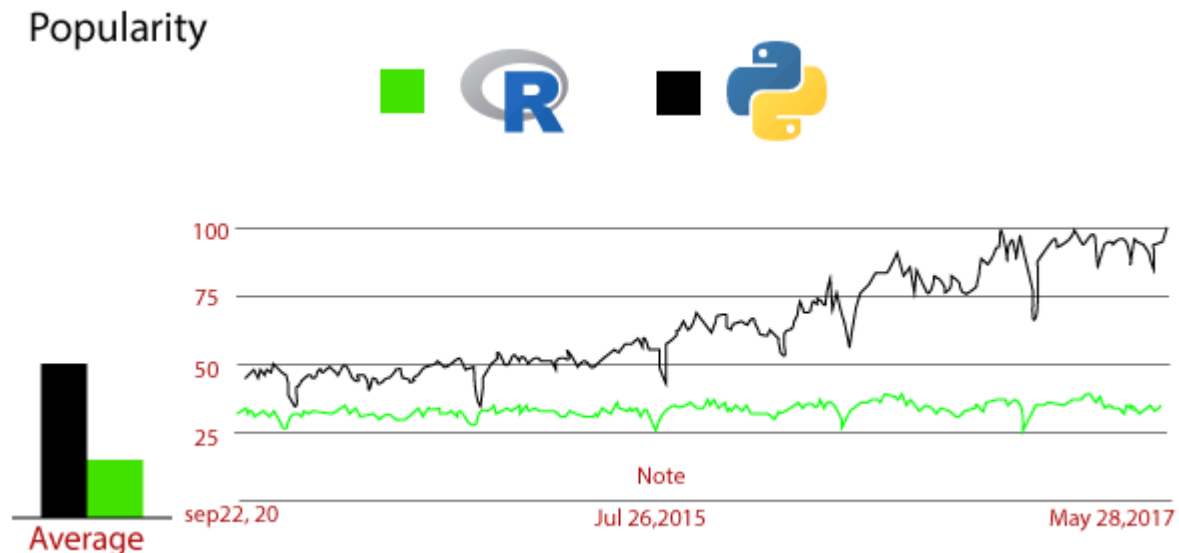
Winner



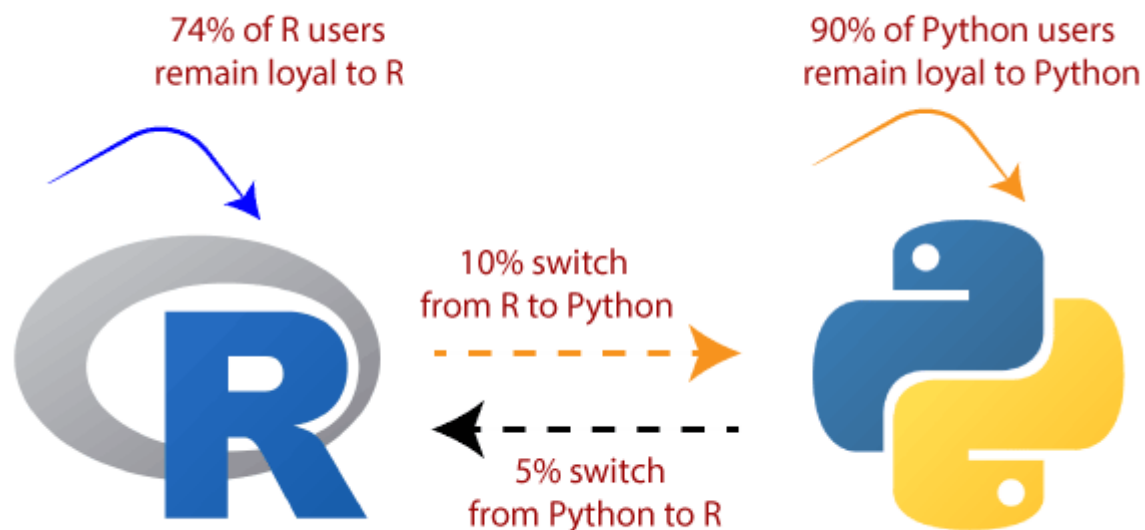
Winner

Popularity

From the past decades, both R and Python were started at the same level. Python became more popular than R. It ranked first in 2016 as compared to R that was ranked 6th on the list. The users of Python are more patriotic rather than R. The percentage of switching from R to Python is twice as large as Python to R.



From the above discussion, we conclude that both languages give a head to head fight in the world of data analytics and data science. But Python emerged as the winner from both due to its widespread popularity and simplicity in writing code.



Difference between R and Python

S.No	Parameters	R	Python
1.	Objective	Its main objective is to perform data analysis and statistics.	Python is used for deployment and production.
2.	Primary Users	Scholar and R&D are the primary users of R.	Programmers and developers are the primary users of R.
3.	Flexibility	In R, we can easily use the available libraries.	In Python, we can easily construct new models from scratch.
4.	Learning Curve	In R, the learning curve is difficult at the beginning.	The learning curve is linear and smooth.
5.	Popularity of programming language	4.23% in 2018.	21.69% in 2018.
6.	Average Salary	\$99.000	\$100.000
7.	Integration	R runs locally.	It is well-integrated with the app.
8.	Task	In R, we can easily get the primary results.	Python is good to deploy the algorithm.
9.	Database size	R handles the huge size of data.	It will also handle the huge size of data.
10.	IDE	Rstudio	Spyder, Ipython Notebook.
11.	Packages and library	tydiverse, ggplot2, caret, and zoo.	Pandas, scipy, scikit-learn, TensorFlow, caret.
12.	Advantages	<ol style="list-style-type: none"> 1. Beautiful graph construction. 2. A large catalog is available for data analysis. 3. It has a GitHub interface. 	<ol style="list-style-type: none"> 1. Notebooks help to share data with colleagues, and python has a jupyter notebook. 2. Mathematical computation. 3. Code Readability.

		4. RMarkdown. 5. Shiny.	4. Function in Python. 5. Deployment. 6. Speed.
13.	Disadvantages	1. It has a slow high learning curve dependencies between the library.	1. Not as many libraries as